DOCUMENT RESUME

ED 306 245 TM 013 014

AUTHOR Myers, Christopher J.
TITLE A Cafeteria Style Approach for Evaluating K-3 Programs.
PUB DATE Mar 89
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Elementary Schools; Elementary School Students; *Evaluation Methods; Evaluation Utilization; Grade 1; Grade 2; Grade 3; Kindergarten; *Primary Education; *Program Evaluation; Research Methodology; Urban Schools

ABSTRACT

A cafeteria-style approach for evaluating primary grade programs in a large metropolitan school district is described. This evaluation approach involved planning between one and four evaluation options for each of 17 program options available to the district's 70 elementary schools. About 20,000 kindergarten through grade 3 students were enrolled in these elementary schools, which ranged in size from 250 to 750 students. From the program and evaluation "menus," school staff could select the combination of programs and evaluation plans that best suited their needs. The multiple evaluation methods used included: (1) multiple measures; (2) multiple designs; (3) multiple analyses; (4) combined qualitative and quantitative methods; and (5) combined multiple evaluation goals. In addition, client participation was an integral part of the process. This approach was well received by school staff because it enabled them to choose options sensitive to the needs of their students. Evaluators found that the approach enhanced internal and external validity, helped explain differences among schools and groups, and enhanced the credibility of findings among evaluation users. Three tables of options are appended. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED306245

A Cafeteria Style Approach for Evaluating K-3 Programs

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Christopher J. Myers

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Christopher J. Myers

Tucson Unified School District
1010 E. 10th St.
Tucson, AZ 85719

Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA, March 27-31, 1989.

TM013014

BEST COPY AVAILABLE

A Cafeteria Style Approach for Evaluating K-3 Programs

Abstract

The purpose of this paper is to describe a cafeteria style approach for evaluating primary grade programs in a large metropolitan school district. The cafeteria style evaluation approach entailed planning between one and four evaluation options for each of 17 program options available to the district's elementary schools. From these program and evaluation "menus", the school staffs were able to select the combination of programs and evaluation plans that would best meet the needs of their students.

At the heart of the cafeteria style evaluation approach was the use of multiple evaluation methods, including multiple measures, multiple designs, multiple analyses, combining qualitative and quantitative methods, and combining multiple evaluation goals. In addition, client participation was an integral part of the process.

From the school staff point of view, the cafeteria style approach was well received because it enabled them to choose options that were most sensitive to the needs of their students. From the evaluator's perspective, this approach enhanced internal and external validity, helped explain differences among schools and groups, and enhanced the credibility of findings among evaluation users.

A Cafeteria Style Approach for Evaluating K-3 Programs

The purpose of this paper is to describe a cafeteria style approach to program evaluation that is being used to evaluate special primary grade programs in a large metropolitan school district. Such an approach may prove valuable to others conducting similar evaluation studies.

The use of multiple and varied evaluation methods has long been advocated for use in evaluating educational programs. While there are some drawbacks, "approaching evaluation from a multimethodological stance . . . allows evaluators to overcome many of the research problems associated with the constraints either imposed by program administrators or inherent in one particular methodology" (Heath, Kendzierski, & Borgida, 1982, p. 234).

The majority of these multiple methods may be categorized into one of the following five classifications:

1. Multiple Measures. Measuring a construct using different instruments and multiple data sources is supported by many researchers within the evaluation community (Cronbach, 1980; Hunter, 1987; Metfessel & Michael, 1973; Reichardt, 1979). One of the reasons for using multiple measures is that the evaluator often is not sure what constitutes the "correct" measure (Cook, 1985). Examples of multiple measures include: (a) measuring the same variable with different instruments; (b) measuring the same variable repeatedly; (c) using a battery of interrelated measures; (d) aggregating observed measures; and (e) measuring different levels, such as the client, group and organization (Lipsey, Cordray & Berger, 1981).

2. Multiple Designs. Through using several different evaluation designs, the evaluator is better able to adapt the evaluation to different sites, times and groups (Lipsey et al., 1981).

3. Multiple Analyses. Multiple data analyses may include (a) conducting different analyses of the same outcome data, or (b) data stratification, which consists of the disaggregation of evaluation data to probe for inconsistencies (Lipsey et al., 1981). Analyzing evaluation data in different ways may help rule out potential biases (Reichardt, 1979).

The author would like to express appreciation to all of the individuals in the Departments of K-3 Programs and Research and Evaluation for their invaluable assistance on this project. In particular, the author would like to thank Rebecca Montano for her leadership and support.

4. Combining Qualitative and Quantitative Methods. Program evaluations may be enhanced through combining traditional quantitative methods with qualitative procedures, such as using open-ended questionnaires or qualitative approaches, such as ethnography (Kidder & Fine, 1987).

5. Combining Multiple Evaluation Goals. Combining different types of evaluation, such as process and outcome, is another way that multiple methods may be used (Mark & Shotland, 1987).

The reasons that multiple evaluation methods are so heartily endorsed by the evaluation community are many. The benefits to using multiple methods include:

1. Enhanced Validity. Since any one measure is fallible, an evaluation's validity may be enhanced through utilizing multiple measures of a variable (Joint Committee on Standards for Educational Evaluation, 1981). Triangulation, which specifies that multiple independent methods be used to converge upon the single best measure of a construct, is an example of a widely used multiple methods approach that helps to rule out threats to validity (Campbell & Fiske, 1959; Cook, 1985; Lipsey et al., 1981; Mark & Shotland, 1987) and reduce errors of measurement (Hunter, 1987).

2. Increased Interpretability of Results. Mark and Shotland (1987) contend that multiple methods allow the evaluator to enhance the interpretability of results through addressing different but complementary evaluation questions. For example, quantitative results may support the effectiveness of a program, while qualitative evaluation findings might be used to explain why the program is effective.

3. Reduced Inappropriate Certainty. When several analyses arrive at the same conclusion, the evaluator is likely to have increased confidence in the findings (Heath et al., 1982), and the findings will have more credibility among consumers of evaluation (Reichardt, 1979).

4. Increased Versatility. Heterogeneous programs and populations should not always be evaluated in the same way because different evaluation approaches may need to be applied in different situations (Patton, 1981). Lipsey et al. (1981) point out that using multiple methods allows for ". . . tailormaking the data collection to fit the exigencies of each site, client [and] group" (p. 285).

Apart from the advantages of using multiple evaluation methods, Patton (1978) has stressed the importance of involving users of evaluation information in the evaluation process. User involvement enhances the users' belief in the data and confidence in the results.

Methods

Several years ago, the Arizona legislature allocated additional funds to be used to enhance instruction for kindergarten through third grade students. Some of these funds are being used by educators in one of the state's largest school districts to implement a myriad of programs, including: teacher inservice training and visitation of other classes; instructional aides; parent training; resource consultants; computer assisted instruction; extended instruction; special writing, literature and math programs; student counseling; and site administrator training.

The evaluator for K-3 Programs was given the task of developing an appropriate plan to evaluate the effectiveness of these varied programs. In addition to considering the different programs needing to be evaluated, the evaluator needed to consider the heterogeneity of the students that would be participating in these programs. Approximately 20,000 K-3 grade students are enrolled in 70 elementary schools ranging in size from 250 to 750 students. Forty-seven percent of the students are from ethnic minorities, with the greatest proportion being Hispanic. The students' home environments range from impoverished to affluent and from dysfunctional to nurturing.

To best meet these evaluation requirements, a cafeteria style evaluation approach was developed. This approach not only utilized multiple evaluation methods, but involved many of the users of the evaluation results in the evaluation process.

The cafeteria style evaluation approach entailed planning between one and four evaluation options for each of the 17 program options available to the schools. The program components and applicable evaluation options were then presented to the staffs at each of the district's 70 elementary schools so that they could select the combination of programs and evaluation plans that would best meet their needs. Tables 1, 2 and 3 (see Appendix) depict program options and corresponding evaluation options that were available to evaluate each program option.

The following multiple methods were used in implementing the cafeteria style evaluation approach:

1. Multiple Measures. Because of the diversity of program components that were implemented, numerous constructs were assessed using a multitude of measures. In some instances, several measures were used to evaluate a construct (e.g., reading achievement), while in other cases only one measure was used.

   Reading and math achievement were evaluated using a nationally normed, standardized achievement test, a district-developed criterion referenced test, textbook unit tests and report card

grades. Writing composition was evaluated by using a holistic writing assessment procedure to rate writing samples.

Other dependent variables were also measured in several different ways. The effectiveness of inservice staff training and parent training was evaluated with subjective rating instruments, while student self-concept was assessed with standardized self-concept measures. In addition, student behavior was evaluated with parent and teacher ratings, frequency counts and report card deportment grades.

Finally, evaluation measures were gathered for several different levels, including classroom, grade level, school and district.

2. Multiple Designs. Several different designs were employed in evaluating the various program components. In some instances, only one design was used with a specific program option, while in other cases, several designs were used. These evaluation designs included (a) nonequivalent control group, pretest-posttest; (b) one group, pretest-posttest; and one group, postassessment.

3. Multiple Analyses. The analyses used for each program option generally corresponded to the evaluation design that was implemented. For interval level data, analysis of variance, repeated measures analysis of variance, analysis of covariance, and correlational analyses were conducted. Nominal level data were generally categorized and summarized. Several analyses were conducted with the same data for some program components, while for other program options a different analysis were used for each set of data.

4. Combining Qualitative and Quantitative Methods. Although ethnography and other qualitative approaches were not used in this evaluation, some qualitative methods were employed. Open-ended questionnaires and interviews were utilized in numerous instances to gather information from principals, teachers, parents and students.

5. Combining Multiple Evaluation Goals. The majority of the evaluation effort was directed toward outcome evaluation so that the evaluation results could be used by decisionmakers in the state department of education, the district central administration office and the schools. However, process evaluation was used on occasion. For example, part of the evaluation of the inservice training program option assessed the degree to which teachers who participated in the training implemented some of the things that they learned from the inservices in their classrooms.

The following example illustrates how the cafeterial style approach works. The principal, I. M. Happy, and the faculty at

Learnalot Elementary School decided to select the Instructional Aides, Inservice Training and Writing and Thinking program options. From among the four available options to evaluate the Instructional Aides program, the Learnalot staff chose Evaluation Option E. For the Instructional Aides program option, paraprofessionals were hired to provide additional academic assistance to at-risk students in the classroom. Evaluation Option B involved having equal numbers of program participants and non-participants tested in the fall and spring with the district-developed criterion referenced reading and/or math test. These data were analyzed using analysis of covariance. In addition, the amount of time that the instructional aides tutored the at-risk students was correlated with the amount of achievement gain between pretesting and posttesting.

The Inservice Training program option, which consisted of having consultants present new ideas and information to the Learnalot staff, was evaluated with Evaluation Option A. This involved having the inservice training participants write a brief proposal specifying how they would implement ideas that were acquired during the training in the classrooms. Principal Happy documented the degree to which the proposals were implemented to determine if 90% of the proposals were successfully implemented.

Writing and Thinking is a commerically produced writing composition program, which was evaluated with Evaluation Option I. Fall and spring writing samples were obtained from a sample of program participants and non-participants. A holistic writing assessment procedure was used to score the samples. Gain scores for the two groups were analyzed using analysis of variance.

Results

Based upon the results of a survey of school principals, the cafeteria style evaluation approach was well received. They liked having the opportunity to make choices among program and evaluation options. The cafeteria style approach enabled the principals and their faculties to select options that would be sensitive to the needs of their students and that would not force them into a "one size fits all" approach.

From the evaluator's perspective, there were both advantages and disadvantages to the cafeteria style approach. On the plus side, use of multiple data sources allowed for a built-in concurrent validation process. That is, findings from one subset of data using a particular instrument could be compared with the results from another subset of data using a different assessment measure. For example, Iowa Test of Basic Skills data were compared with district criterion referenced test data for the same students as well as students from different schools to evaluate the instructional aides program. In those instances where there was agreement among results, the evaluator could be particularly confident in conclusions drawn from the findings.

Secondly, by conducting analyses across several different levels, it was possible to determine which programs were most effective at specific schools, at specific grade levels and for the district as a whole, thus enhancing external validity. It was also possible to identify the programs that were most effective with students having different demographic characteristics.

Finally, the use of qualitative measures enabled the evaluator to gather data that could not be obtained through quantitative means. While the outcome evaluation results helped the evaluator assess which schools effectively implemented successful programs, qualitative data helped to decipher why programs were successful at some schools but not at others.

On the minus side, considerably more time was required for the evaluator to develop the evaluation plan, and to collect and analyze the data. However, the benefits derived from this approach were felt to offset the disadvantages.

Conclusions

The cafeteria style evaluation approach appears to have merit in situations where multiple programs are being implemented in different settings. Specifically, this approach provides the following benefits:

1. It allows for "client" participation by enabling evaluation consumers to choose evaluation options, thus enhancing credibility of the findings.

2. In those instances when different program components are being evaluated using the same measures, it allows for interprogram comparisons to be made.

3. In those instances when the same program component is being evaluated with different assessment measures or procedures, it allows for validation of the findings.

4. It provides flexibility, allowing for recognition of differences among student populations and school sites.

5. It allows the evaluator to determine which program and evaluation options appear to work best under which conditions.

References

Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validations by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Cook, T. D. (1985). Postpositivist critical multiplism. In R. L. Shotland & M. M. Mark (Eds.), Social Science and Social Policy. Beverly Hills: Sage Publications.

Cronbach, L. J. (1980). Toward reform of program evaluation. San Francisco: Jossey-Bass.

Heath, L., Kendzierski, D. & Borgida, E. (1982). Evaluation of social programs: A multimethodological approach combining a delayed treatment true experiment and multiple time series. Evaluation Review, 6, 233-246.

Hunter, J. E. (1987). Multiple dependent variables in program evaluation. In M. M. Mark & R. L. Shotland (Eds.), Multiple methods in program evaluation [monograph]. New Directions for Program Evaluation, 35, 43-56.

Joint Committee on Standards for Educational Evaluation (1981). Standards for Evaluations of Educational Programs, Projects and Materials. NY: McGraw-Hill Book Co.

Kidder, L. H. & Fine, M. (1987). Qualitative and quantitative methods: When stories converge. In M. M. Mark & R. L. Shotland (Eds.), Multiple methods in program evaluation [monograph]. New Directions for Program Evaluation, 35, 57-75.

Lipsey, M. W., Cordray, D. S. & Berger, D. E. (1981). Evaluation of a juvenile diversion program: Using multiple lines of evidence. Evaluation Review, 5, 283-306.

Mark, M. M. & Shotland, R. L. (1987). Alternative models for the use of multiple methods. In M. M. Mark & R. L. Shotland (Eds.), Multiple methods in program evaluation [monograph]. New Directions for Program Evaluation, 35, 95-100.

Metfessel, N. S. & Michael, W. B. (1973). A paradigm involving multiple criterion measures for the evaluation of the effectiveness of school programs. In B. R. Worthen & J. R. Sanders (Eds.), Educational evaluation: Theory and practice (pp. 269-289). Bellmot, CA: Wadsworth Publishing Co.

Patton, M. Q. (1978). Utilization-focused evaluation. Beverly Hills: Sage Publications.

Patton, M. Q. (1981). Creative evaluation. Beverly Hills: Sage Publications.

Reichardt, C. S. (1979). The statistical analysis of data from nonequivalent group designs. In T. D. Cook & D. T. Campbell (Eds.), Quasi-experimentation: Design and analysis issues for field settings (pp. 147-205). Skokie, Il: Rand McNally.

APPENDIX

Table 1

1988-89 K-3 Program and Evaluation Options

Program Options		Evaluation Options											
		A	B	C	D	E	F	G	H	I	J	K	L
1.1	Inservice Training	X											
1.2	Visitation of Other Classes and Programs	X											
1.3	Instructional Aides		X	X	X	X							
1.4	Project 3T						X						
1.5	Parent Training							X					
1.6	Resource Consultant		X	X	X	X			X				
1.7	Computer Assisted Instruction (CAI)		X	X	X	X							
1.8	Extended Instruction-After Hours Tutoring		X	X	X	X							
1.9	Extended Instruction-Project HIP		X	X		X							
1.10	Extended Instruction - Kinder Bridge											X	
1.11	Writing and Thinking									X			
1.12	Second or Third Grade Literature Study		X	X	X	X							
1.13	Daily Oral Math (DOM)		X	X	X	X							
1.14	Student Advocate										X		
1.15	Administrative Support-Supervision Workshop												X
1.16	Administrative Support-Peer Supervision												X
1.17	Brigance Screening												X

Table 2

1988-89 K-3 Program Options

	Program Option	Description
1.1	Inservice Training	Selected staff participated in training designed to enhance knowledge and skills that would directly impact at-risk students.
1.2	Visitation of Other Classes and Programs	Selected staff observed classrooms of other teachers or other educational programs to acquire skills that could be used with at-risk students.
1.3	Instructional Aides	Paraprofessionals assisted classroom teachers by tutoring at-risk students or monitoring the work of other students so the teacher could work directly with at-risk students.
1.4	Project 3T (Teachers Training Teachers)/ TOPS (Think, Originate, Plan and Solve)	Trainers modeled techniques for teaching higher order thinking skills alternated with the trainers observing and critiquing the trainees' skills.
1.5	Parent Training	Parents of at-risk students participated in training to help them better understand their children or to assist their children with academic tasks at home.
1.6	Resource Consultant	A content area specialist instructed at-risk students in health, language arts, mathematics, music, science, social studies or visual arts.
1.7	Computer Assisted Instruction (CAI)	At-risk students received CAI in reading and/or math to teach new skills or to reinforce previously taught skills.
1.8 1.9 1.10	After Hours Tutoring Project HIP (Homework In Progress) Kinder Bridge	Variations on providing tutoring to at-risk students after regular school hours for the purpose of enhancing their academic skills by increasing the amount of instruction received.
1.11	Writing and Thinking	Writing composition program that teaches writing as a process using six phases.

continued

TABLE 2 (contd.)

Program Option		Description
1.12	Literature Study	Supplemental literature program designed to enhance reading comprehension among second and third graders.
1.13	Daily Oral Math (DOM)	Materials designed to provide daily practice and reinforcement of math skills for students in Grades K-3.
1.14	Student Advocate	Selected school staff were given release time to counsel students who are experiencing social or emotional problems.
1.15	Supervision Workshops	Training was provided selected school administrators in key administrative areas.
1.16	Peer Supervision	Selected staff were trained in a collaborative model of supervision that enabled them to assist selected teachers to improve instruction.
1.17	Brigance Screening	Selected teachers were trained to use the Brigance assessment and intervention program.

Table 3

1988-89 K-3 Evaluation Options

Eval. Option	Evaluation Procedures/ Instruments/Sample	Analysis
A	Participants completed a brief proposal describing how they would utilize newly-acquired skills with at-risk students.	The principal monitored proposal implementation and determined if 90% of participants successfully implemented proposed activities
B	Equal numbers of target students* and control students* were tested in the fall and spring with district criterion referenced reading and/or math tests (SOS). Logs were kept of the type and amount of assistance provided.	Analysis of variance (ANOVA), analysis of covariance (ANCOVA), and/or correlation between achievement gain and amount of tutoring
C	Target students* were tested in the fall and spring with reading and/or math SOS. Logs were kept of the type and amount of assistance provided.	ANOVA, ANCOVA and/or correlation between achievement gain and amount of tutoring
D	Target students* were tested in the springs of consecutive years with the Iowa Test of Basic Skills (ITBS). Logs were kept of the type and amount of assistance provided.	ANOVA, ANCOVA and/or correlation between achievement gain and amount of tutoring
E	Target students* were tested in the fall and spring with the Addison-Wesley math and/or Scott-Foresman reading tests. Logs were kept of the type and amount of assistance provided.	ANOVA, ANCOVA and/or correlation between achievement gain and amount of tutoring
F	Samples of target students* and control students* were interviewed about TOPS and trainees completed questionnaires before and after Project 3T/TOPS was implemented.	ANOVA for quantitative data and descriptive statistics for qualitative data
G	Participating parents completed assessment forms following training.	Descriptive statistics and summary of narrative comments

continued

TABLE 3 (contd.)

Eval. Option	Evaluation Procedures/ Instruments/Sample	Analysis
H	Equal numbers of target students* and control students* were tested before and after program implementation with a content appropriate achievement measure.	ANOVA and/or ANCOVA
I	Target students and control students produced fall and spring writing samples that were holistically scored.	ANOVA and/or ANCOVA
J	Attendance, disciplinary referrals/actions and deportment grade data were collected for target students* before and after program implementation. Logs were kept of number and types of interventions.	ANOVA for quantitative data and descriptive statistics for qualitative data
K	Equal numbers of target students* and control students* were tested in the fall and spring with the Battelle Developmental Screening Test.	ANOVA and/or ANCOVA
L	Participants completed assessment forms following training or assistance.	Descriptive statistics and summary of narrative comments

*Target students participated in the specified program; control students did not participate in the specified program.